Assimilating Along-Track Altimetric Observations Through Local Hydrostatic Adjustment in a Global Ocean Variational Assimilation System

Andrea Storto* Srdjan Dobricic

Centro Euro-Mediterraneo per i Cambiamenti Climatici, Bologna, Italy

Simona Masina

Istituto Nazionale di Geofisica e Vulcanologia, Bologna, Italy and

Centro Euro-Mediterraneo per i Cambiamenti Climatici, Bologna, Italy

Pierluigi Di Pietro

Istituto Nazionale di Geofisica e Vulcanologia, Bologna, Italy

* *Corresponding author address:*, Andrea Storto, Centro Euro-Mediterraneo per i Cambiamenti Climatici, viale Aldo Moro 44, I-40127 Bologna, Italy.

E-mail: andrea.storto@cmccc.it

ABSTRACT

A global ocean three-dimensional variational data assimilation system was developed with the aim of assimilating along-track sea-level anomaly observations, along with in-situ observations from bathythermographs and conventional sea stations. All the available altimetric data within the period October 1992 to January 2006 were used in this study. The sea-level corrections were covariated with vertical profiles of temperature and salinity according to the bivariate definition of the background-error vertical covariances. Sea-level anomaly observational error variance was carefully defined as a sum of instrumental, representativeness, observation operator and mean dynamic topography error variances. The mean dynamic topography was computed from the model long-term mean sea-surface height and adjusted through an optimal interpolation scheme to account for observation minus first guess biases. Results show that the assimilation of sea-level anomaly observations improves the model sea-surface height skill scores as well as the subsurface temperature and salinity fields. Furthermore, estimate of the tropical and sub-tropical surface circulation is clearly improved after assimilating altimetric data. We have also found non-negligible impact of the mean dynamic topography used: with respect to a gravimeter-based mean dynamic topography the use of our mean dynamic topography improves both the consistency with sea-level anomaly observations and the verification skill scores of temperature and salinity in the Tropical regions. Furthermore, the use of a mean dynamic topography computed from the model long-term sea-surface height mean without observation adjustments results in worsened verification skill scores and highlights the benefits of our approach for deriving the mean dynamic topography.

1. Introduction

The study of the global ocean inter-annual and lower frequency variability requires the accurate knowledge of the surface and subsurface ocean state. This is usually obtained by optimally combining in-situ and satellite observations. The latter are useful especially in those areas where the network of conventional observations is poor (e.g. in the Extra-Tropical South Hemisphere).

In the last three decades, many altimetric satellite missions have been launched with the aim of monitoring variations of sea-level height, both at regional and global scale. Sea-level height measured by space-borne altimeters refer to a reference ellipsoid and not to the spatially-varying elevation of the Earth geoid. Therefore, the measurements are provided in terms of anomalies from the time-averaged sea-level height referenced to the Earth geoid, which is usually called Mean Dynamic Topography (MDT). The information brought by the sea-level anomaly data (SLA), in terms of temporal and spatial variability of sea-level height, has been successfully exploited within data assimilation systems to adjust surface and subsurface temperature and salinity fields. For instance, Zheng et al. (2007) demonstrated that assimilating SLA observations improves the predictions of sea surface temperature (SST) anomaly in the Equatorial Pacific. In other works (Fischer et al. 1997; Ji et al. 2000), the prediction of ENSO in seasonal forecasts is improved by using the sea-level information only to correct the first corresponding subsurface temperature levels. However, evidence of local importance of the halosteric contribution to the sea-surface height trend (Ivchenko et al. 2008) suggests that sea-level anomaly data can be fruitfully assimilated to correct also subsurface salinity fields. This was also confirmed by the positive impact of the bivariate

assimilation of altimetric data found by Vossepoel and Behringer (2000), although they used synthetic observations.

The definition of a global Mean Dynamic Topography, to be added to the sea-level anomaly observations, is required for the assimilation of sea-level anomalies in order to compare the observations with the model-equivalents. This is a non-trivial issue if one considers that even though the very precise recent gravimetric data available from the the Challenging Minisatellite Payload (CHAMP), from the Gravity Recovery And Climate Experiment (GRACE) or even from the higher-resolution just launched Gravity Field and Steady-State Ocean Circulation Explorer (GOCE) can provide a precise estimate of the geoid, the ocean model represents the geoid in a much more simplified way. Moreover, even with a precise estimate of the geoid, it would need very accurate sea-ice, river runoff and steric variability parametrizations to correctly estimate the absolute value of the sea-surface height and correctly follow its seasonal and inter-annual variations.

As explained in details by Vidard et al. (2009), an observed MDT, such as the one from gravimetric missions, might be improper as it bears a geodetic bias which varies locally and can not simplify to a global offset. This is even more dramatic since the impact of the choice of the MDT can be non-negligible. Segschneider et al. (2000), for instance, showed the sensitivity of the ocean forecasts to the MDT, which can induce variations in the Niño-3 averaged 100-meter temperature up to 4°Celsius. As a matter of course, the importance of a correctly estimated MDT and its consistence with the model sea-surface height is a key-issue for the successful assimilation of altimetric data.

At the Euro-Mediterranean Centre for Climate Change (CMCC) a three-dimensional variational (3DVAR) data assimilation system called OceanVar has been developed by Do-

bricic and Pinardi (2008) in the framework of the Mediterranean Forecasting System (MFS) (Pinardi et al. 2003), producing operationally daily analysis and forecasts (Dobricic et al. 2008). On the other hand, a global ocean analysis system was implemented for reanalysis and climate research applications, see for example Masina et al. (2010); Lee et al. (2010). This assimilation system (Bellucci et al. 2007) was the global implementation of a reduced-order multi-variate Optimal Interpolation (OI) scheme (De Mey and Benkiran 2002) and all the available in-situ observations were assimilated. Although the resolution of the analysis system is rather coarse (i.e. 2 degrees with increase up to 0.5 degrees at the Equator), such a configuration is still very attractive and necessary for many applications like long-term climate studies (Masina et al. 2010) and seasonal ensemble prediction system (Alessandri et al. 2010), where the use of higher resolutions is still too demanding in terms of computer resources.

As the CPU time and memory demand of an OI scheme primary relies on the amount of observations, it had the limitations of hardly being suitable for the assimilation of many space-borne instruments which enormously increase the length of the observation vector. On the contrary, 3DVAR computing resources depend on the dimensions of the computational grid. Note also that an increase of resolution leads not only to an increase of the model state vector but also to an increase of observation vector, as the observation selection strategies (e.g. subsampling, thinning, etc.) strictly rely on the grid resolution. Moreover, 3DVAR provides global solutions of the analysis problem without leading to large and small scale inconsistencies as in OI (Lorenc 1981), and allows for the implementation of complex observation operators in a more efficient way (see e.g. Kalnay 2002, Section 5.5.3). In order to produce global analyses with the possibility of assimilating space-borne altimetric data

and, in the future, sea-surface temperature and salinity measurements from satellite, we have adapted the limited-area OceanVar to the global ocean and implemented a scheme for locally adjusting the water column-integrated density through the sea-level anomaly observations.

This paper is concerned with the description and the validation of the SLA data assimilation. The structure is as follows: Section 2 briefly describes the 3DVAR/FGAT formulation and the configuration of the analysis system; in Section 3 the implementation of a scheme for assimilating SLA observations is presented, together with the procedure and the results of the computation of the global Mean Dynamic Topography (MDT); Section 4 presents the results obtained with the assimilation of altimetric data, while Section 5 summarizes the main achievements and proposes a discussion on the results.

2. The assimilation system

a. 3DVAR/FGAT formulation

Three-dimensional variational assimilation methods iteratively solve the assimilation problem by minimizing a cost function J (Courtier 1997), given by

$$\begin{aligned} J = & \frac{1}{2}\left(\mathbf{x}-\mathbf{x^b}\right)^{\mathsf{T}}\mathbf{B}^{-1}\left(\mathbf{x}-\mathbf{x^b}\right)+ \\ & +\frac{1}{2}\left(\mathbf{y}-\mathbf{H}(\mathbf{x}-\mathbf{x^b})-\mathcal{H}(\mathbf{x^b})\right)^{\mathsf{T}}\mathbf{R}^{-1}\left(\mathbf{y}-\mathbf{H}(\mathbf{x}-\mathbf{x^b})-\mathcal{H}(\mathbf{x^b})\right), \end{aligned} \tag{1}$$

whose minimization requires the computation of its gradient ∇J:

$$\nabla J = \mathbf{B}^{-1}\left(\mathbf{x}-\mathbf{x^b}\right)-\mathbf{H}^{\mathsf{T}}\mathbf{R}^{-1}\left(\mathbf{y}-\mathbf{H}(\mathbf{x}-\mathbf{x^b})-\mathcal{H}(\mathbf{x^b})\right) \tag{2}$$

where $\mathbf{x}$ is the unknown ocean state, equal to the analysis $\mathbf{x^a}$ at the minimum of J, $\mathbf{x^b}$ is the background, which is an *a priori* estimate of the state of the ocean, $\mathbf{y}$ is the vector

of the observations, $\mathcal{H}$ is the fully non-linear *observation operator* which projects the state of the ocean onto the space of the observations (i.e. an interpolation operator from model grid to observation location and eventually a transformation operator from model variable to observation variable), $\mathbf{H}$ is the tangent-linear version of the *observation operator*, $\mathbf{H}^{\mathsf{T}}$ its adjoint version and $\mathbf{B}$ and $\mathbf{R}$ are the covariance matrices of the background and observational errors, respectively. In the formulation of Equation (1), the fully non-linear observation operator is used only once for computing the initial departures using the background fields closer to observation time (the so-called First Guess at Appropriate Time, FGAT). The tangent-linear model is used for updating the cost function at each iteration according to the new model state, while the adjoint model for mapping the new observation departures back onto the model space for the gradient computation. Their linearization is performed around the background fields closer to observation time. In our formulation, only temperature and salinity are corrected after a 3DVAR/FGAT assimilation step, so that the model space is composed of the pair of (T, S) in the three-dimensional ocean grid. The cost function minimization is numerically achieved through the quasi-Newton L-BFGS minimizer of Byrd et al. (1995).

The formulation of the background term of the cost function follows Dobricic and Pinardi (2008): the background-error covariance matrix is decomposed in a sequence of linear operators which account separately for horizontal correlations and vertical covariances, assumed to be independent. Horizontal correlations are modeled by means of an application of four iterations of a first-order recursive filter (Hayden and Purser 1995) with horizontally homogeneous but vertically varying correlation lengthscale.

The vertical component of the background-error covariance matrix is eigendecomposed to

obtain vertically uncorrelated eigenvectors. We use empirical orthogonal functions (EOFs), which are bivariate in temperature and salinity, at full model resolution on both the horizontal and the vertical. In order to reduce computational cost and avoid noisy vertical correlations, the vertical modes are truncated to 10 from the original 62 modes. The EOFs were calculated from the model seasonal climatology (Bellucci et al. 2007; Di Pietro and Masina 2009).

In order to impose cyclic conditions on the global ocean domain, the minimization is performed over an *extended domain*, with duplicated points before the Western and after the Eastern boundaries. Observations in the extension zones are duplicated from the corresponding area of the original domain. The extension zone is 20 grid-points long on both boundaries (about 5000 Km), which makes sure that analyzed fields at the inner boundaries of the extension zones are exactly the same at the Western and Eastern boundaries after the minimization.

The data assimilation step is performed every 10 days, using all the observations included in the temporal range of ± 5 days before and after the assimilation time. The ocean model is then used to project the analyzed fields forward to the next assimilation step.

b. In-situ observations

The set of in-situ observations consists of vertical profiles of temperature from the Expandable bathythermographs (XBT), buoys, sea stations (TESAC), Argo floats (from late 90s onwards only) and salinity profiles from buoys, sea stations and Argo floats. Data have been provided by the U.K. Met Office Hadley Centre in the framework of EU-funded project

ENSEMBLES (EN3). EN3 is a collection of in-situ data which have been quality-checked according to the method of Ingleby and Huddleston (2007). For XBT data, a time-dependent fall rate correction was used (Wijffels et al. 2008).

An additional background quality check, which rejects observations whose square departure from the model-equivalents is larger than three times the sum of the background and observational error variances is also performed on top of the minimization to avoid inconsistencies between the model and the observations. Finally, a thinning procedure retains only the closest observation to analysis time in case that multiple reports from the same platform or buoy are found in close proximity.

c. The ocean model

The oceanic general circulation model (OGCM) used within the analysis system is the free-surface Ocean Parallèle model (OPA) in its version 8.2 (Madec et al. 1998). The spatial resolution of OPA is $2°$x$2° \cos(latitude)$x31 vertical levels, though in the tropical band the meridional resolution increases up to $0.5°$.

Surface boundary conditions are provided by the European Center for Medium-range Weather Forecasts (ECMWF) ERA-40 reanalysis dataset (Uppala et al. 2005) for wind stress, surface heat and freshwater fluxes from 1990 till 2001. Afterwords, they are provided by the ECMWF operational analyses; this leads to non-continuous forcing fields at the beginning of 2002 and to a possible deterioration of the quality of the forcing fields, but we kept this setup for comparison with the previous analysis system (Masina et al. 2010). A relaxation to the Reynolds SST (Reynolds and Smith 1994), and to the ECMWF operational

SST from 2001 onwards for consistency with the wind stress and the heat fluxes, was used with a restoring time-scale of about 12 days. Furthermore, our model configuration does not allow any change in the global ocean volume, which is conserved throughout the simulations by superimposing on a daily basis zeroed variation in the globally-averaged water flux. This constraint has been used to avoid unrealistically very large positive trends in the sea-level which were observed without the superimposed closure of the freshwater budget. To better appreciate the impact of sea-level anomaly observations, we do not use any relaxation to climatological fields, as they may too strongly constrain the model temperature and salinity forecasts. The first 3DVAR run was initialized by using the OI analysis fields from Di Pietro and Masina (2009) as initial first guess.

3. Assimilation of SLA observations

a. Sea-level anomaly dataset

Sea-level anomaly observations used in this work covered the period between October 1992 and January 2006, during which many altimetric missions have been providing sea-level anomaly data. Table 2 summarizes the temporal coverage of the altimetric satellites, which have all been assimilated in the 3DVAR/FGAT assimilation system.

Along-track data were provided by AVISO, after the usual geophysical removals (tropospheric, ionospheric, electromagnetic, tidal and inverse barometer effects) and after a multi-satellite cross-calibration for eliminating residual orbit errors (Le Traon et al. 1995) and large-scale biases (Le Traon and Ogor 1998; Le Traon et al. 1998). The dataset of

along-track data is the one provided in delayed time, which has a better orbital calculation, a centered time-window for the multi-satellite cross-calibration and was proved to have a significantly better accuracy (Pascual et al. 2009).

To filter out small spatial scales which are not resolved by the model, we apply a one-dimensional along-track low-pass Lanczos filter (Duchon 1979) with a latitudinally-varying cut-off wavelength consistent with the model resolution, which is about 75 Km close to Equator and decreases polewards. The globally-averaged reduction of root mean square error of SLA observation minus first guess due to the implementation of the Lanczos filter (not shown) is constantly of about 2 mm for the 1993-2005 period. Like the in-situ observations, sea-level anomaly observations are quality-checked against model-equivalents data. Finally, a satellite-dependent thinning procedure retains only one active observation over a box of size 150% the model resolution. This ensures the consistency between the resolution of the assimilated SLA dataset and that of the ocean model, and decreases both the horizontal correlation of the SLA observational errors and the along-track correlation introduced by the Lanczos filter. Briefly, data thinning prevents the neglected observational error correlation from causing inaccuracies in the analyses. In the future, we plan to release the assumption of uncorrelated SLA observations, which in turn will allow to assimilate as many observations as possible.

b. The observation operator

The observed sea-level anomaly $\boldsymbol{\zeta^o}$ is directly related to the sea surface height $\boldsymbol{\eta}$ through the relation

$$\boldsymbol{\zeta^o} = \boldsymbol{\eta} - \boldsymbol{\eta_M} + \boldsymbol{\varepsilon} \tag{3}$$

where $\boldsymbol{\eta_M}$ is the Mean Dynamic Topography (MDT), $\boldsymbol{\eta}$ is the sea surface height (SSH) and $\boldsymbol{\varepsilon}$ is the sum of the errors of the SLA observation, the sea surface height and the MDT. According to the 3DVAR/FGAT formulation of Equation (1), the SLA contribution to the observational term of the cost function is

$$\mathbf{y} - \mathbf{H}\left(\mathbf{x} - \mathbf{x^b}\right) - \mathcal{H}\left(\mathbf{x^b}\right) = \boldsymbol{\zeta^o} - \mathbf{H}\left(\mathbf{x} - \mathbf{x^b}\right) - \left(\boldsymbol{\eta^f} - \boldsymbol{\eta_M}\right) \tag{4}$$

where $\boldsymbol{\eta^f}$ is the background sea surface height prognosed by the ocean model. In the case of the linear free surface formulation of the OPA ocean model (Roullet and Madec 2000), we have that

$$\frac{\partial \boldsymbol{\eta^f}}{\partial t} = \frac{\mathbf{E} + \mathbf{R} - \mathbf{P}}{\rho_w} - \boldsymbol{\nabla u}|_{z=\eta}\,, \tag{5}$$

where $\mathbf{E}$, $\mathbf{R}$ and $\mathbf{P}$ are the terms of the surface freshwater budget (evaporation, river runoff and precipitation, respectively), ρ_w is the water density and $\boldsymbol{u}$ the horizontal current vector. Equation (5) is computed at each model timestep. Within the 3DVAR/FGAT assimilation system, the computation relative to the timestep closest to observation time determines the initial observation minus first guess departure.

To compute the tangent-linear increments of the SLA model-equivalents within the minimization, we use a "local hydrostatic adjustment" (LHA) scheme, which is based on the

vertical integration of density increments:

$$\rho_0 g \delta\eta + \int_{-h_b}^{0} g\delta\rho(T,S)\mathrm{d}z = \delta p_b \tag{6}$$

where $\delta\eta$ is the sea-level height increment and the density increment $\delta\rho$ is integrated over the vertical between the bottom depth h_b and the surface. The $\rho(T,S)$ relation is calculated by means of the 1980 UNESCO International Equation of State (IES 80) as described in Fofonoff (1985). We assume the existence of a "level of no-motion", corresponding to a depth h^* where horizontal velocities are practically zero. This implies, through geostrophy, that the corresponding pressure increment δp_{h^*} vanishes too. Our linearized observation operator then becomes

$$\mathbf{H}\left(\mathbf{x}-\mathbf{x^b}\right) = -\frac{1}{\rho_0}\int_{-h^*}^{0}\delta\rho(\mathbf{x}-\mathbf{x^b})\mathrm{d}z, \tag{7}$$

where we have omitted the horizontal interpolation operator for sake of simplicity. Equation (7) comes from the geostrophic relation, and it is therefore not suitable in close proximity of the Equator. Consequently, SLA observations are not assimilated between 2S and 2N. A similar formulation was proved by Cooper and Haines (1996) and Alves et al. (2001) to successfully correct the water-column integrated density without altering the water properties. We have set the "level of no-motion" to 1500 m of depth, whereas applicable, though the sensitivity of the analyses to this parameter was found very small in the range 1000-2000 m of depth. The sea-level increments obtained through Equation (7) have an accuracy of about 1 to 2 cm with respect to the prognostic formulation of Equation (5) (see Section d), which is below the instrumental accuracy of altimetric data (Le Traon and Ogor 1998).

In practice, our scheme splits the observation departure in thermo- and halo-steric contributions over the water column by using the adjoint version of the density operator linearized

around the background fields of temperature and salinity. How the sea-level increment is vertically spread on the temperature and salinity increments depends upon the bivariate definition of the vertical background-error covariances, which are spatially non-uniform. Note that this approach differs from that previously adopted within OceanVar (Dobricic and Pinardi 2008), where a barotropic model operator was used to simultaneously constrain currents, temperature, salinity and sea-level height fields. The analysis increments structure for the same sea-level anomaly observation minus first guess departure may vary significantly with the observation location. For instance, the vertical background-error signal of the temperature in the Tropical Pacific reaches a maximum at around 150 m of depth (not shown). The sea-level anomaly observation information is spread accordingly and its temperature analysis increments have a maximum at such a depth. On the contrary, vertical background-error variances and sea-level analysis increments are contained in the first 100-150 m of depth within extra-tropical regions, with a maximum typically located in the first 50 m of depth.

Further, the seasonality of the background-error structures is clearly felt by the SLA observation assimilation: North-Atlantic summer-time assimilation of SLA causes the largest analysis increments limited in the first 50 m of depth because of the ocean stratification, while during winter the analysis increments are much more smoothed along the vertical.

c. Global Mean Dynamic Topography

Sea-level anomaly observations require a reference Mean Dynamic Topography to be added to for comparison with the model sea surface height. Purely observational method, which are based on space-borne gravimetric data (Tapley et al. 2003) and eventually ad-

justed by means of in-situ observations (Rio and Hernandez 2004, hereafter RIO04) may not represent the Mean Dynamic Topography seen by the ocean model (Vidard et al. 2009). We have therefore firstly derived the MDT as the model long-term mean SSH from analyses and forecasts initialized by assimilating in-situ data only. Then, we adjust this MDT through output assimilation diagnostics. Indeed, following Dobricic (2005), the different contributions to the error in Equation (3) can be identified. The innovation $\mathbf{d}$ (i.e. the observation minus first guess) is consequently equal to

$$\mathbf{d} = \boldsymbol{\varepsilon}^{\boldsymbol{o}} - \boldsymbol{\varepsilon}^{\boldsymbol{f}} + \boldsymbol{\varepsilon}^{\boldsymbol{M}} \tag{8}$$

where $\boldsymbol{\varepsilon}^{\boldsymbol{o}}$ is the SLA observational error (instrumental and representativeness), $\boldsymbol{\varepsilon}^{\boldsymbol{f}}$ is the error associated to the background field and $\boldsymbol{\varepsilon}^{\boldsymbol{M}}$ is the error associated with the Mean Dynamic Topography. By averaging over several assimilation cycles one obtains

$$\overline{\mathbf{d}} = \overline{\boldsymbol{\varepsilon}^{\boldsymbol{o}}} - \overline{\boldsymbol{\varepsilon}^{\boldsymbol{f}}} + \overline{\boldsymbol{\varepsilon}^{\boldsymbol{M}}}. \tag{9}$$

Assuming that the model and the observations errors are unbiased (or at least their contribution to the innovation bias is unimportant with respect to the MDT bias contribution), it is possible to *a posteriori* correct the bias of the Mean Dynamic Topography by subtracting the bias of the innovation $\overline{\mathbf{d}}$. Dobricic (2005) pointed out that this procedure can be applied as a fixed-point algorithm for successive corrections of the MDT, provided that the MDT contribution to the total bias is still dominating. Unlike (Dobricic 2005), who corrected the MDT by adding the gridpoint SLA bias, we have performed an univariate Optimal Interpolation of the observation departures in order to correct the MDT in practice, according to

which the analyzed Mean Dynamic Topography is equal to:

$$\boldsymbol{\eta}_M^a = \boldsymbol{\eta}_M^b + \frac{\mathbf{B}_\eta}{\mathbf{B}_\eta + \mathbf{R}_\eta}\mathbf{d} \tag{10}$$

where $\boldsymbol{\eta}_M^b$ is the background MDT (i.e. the model long-term mean SSH), $\mathbf{B}_\eta$ is the MDT background-error variance, calculated as model SSH anomaly standard deviation, and $\mathbf{R}_\eta$ the diagonal innovation bias error covariance matrix. The computation is performed by assuming that MDT background-errors are spatially uncorrelated, that is $\mathbf{B}_\eta$ simplifies to an error variance value.

The contribution of each innovation is weighed by the distance between the observation location and the gridpoint through a function of the form

$$c(dx, dy) = \exp\left(-\left(\frac{dx^2}{R_x^2} + \frac{dy^2}{R_y^2}\right)\right), \tag{11}$$

where dx and dy are their zonal and meridional distances and R_x and R_y are the distance scale coefficients. They were set equal to 120 Km, in accordance with the mid-latitude altimetry correlation scales found by Le Traon and Hernandez (1992). The minimum distance for an observation to be included in the innovation vector for a certain gridpoint is 1000 Km. This procedure was performed using all the SLA assimilation statistics within the period 1993-2005, and giving a satellite-dependent weight to the innovation statistics in accordance with the satellite instrumental accuracy.

It is important to stress that our method leads to a MDT (hereafter MDTOI) which is consistent with both the OGCM dynamics and the in-situ and altimetric observations. Therefore it differs from gravimetry-based MDTs, which usually do not account for the ocean model dynamics, needed within data assimilation systems, and use observations mostly to

increase the coarse resolution of remotely sensed gravimetry data. MDTOI differs also from model only-based MDTs, which usually do not account for altimetry data.

MDTOI (Figure 1a) owns very similar features of other mean dynamic topographies in terms of large-scale circulation and strongest gradients in correspondence of main large-scale current systems, western boundary currents and Antarctic Circumpolar Current. We also show the difference between the RIO04 MDT and that computed with our two-step procedure (Figure 1b), and the difference between a model only-based MDT (hereafter MODMDT) and the MDTOI (Figure 1c). For the latter, the MODMDT was derived from a long-term sea surface height mean of a model run without any data assimilation and with the surface boundary forcing as described in Section 2c.

In the comparison with the RIO04 MDT, we found that the main differences are located in the North Atlantic, in correspondence of the Antarctic Circumpolar Current and in areas of large eddy activity (e.g. Kuroshio and Agulhas current systems). The differences cannot be simplified to a global offset, and reach values up to 40 centimeters; they are particularly large and irregular with many complicated features in the mid- and high-latitude southern hemisphere. Note that in these areas, in addition to the presence of many complicated current systems, the lack of a dense network of in-situ observations decreases the accuracy of the RIO04 MDT. Main differences between the model only-based MDT and the MDTOI are in general located in the same areas, although the features are more regular and smoothed. It is important to note that in many dynamically important regions such as the Gulf Stream, the sub-tropical South Atlantic, the Southern Pacific and partly the Kuroshio regions, the differences between the MODMDT and the MDTOI are of opposite sign with respect to the differences between the RIO04 MDT and the MDTOI. This means that the effect of

assimilating in-situ observations and adjusting the MDT with the altimetric observation biases brings our MDT closer to a gravimetry-based one, and leads in turn to an MDT which accounts for the OGCM dynamical features and mitigates its inaccuracies by means of observational data ingestion.

Surface geostrophic currents (not shown) deducted from the RIO04 MDT differ with respect of those deducted from the MDTOI especially in the Gulf Stream Region and in correspondence of the Caribbean, the Falkland, the Agulhas and Equatorial Counter currents, and, partly, in the Kuroshio current area. This can be partly explained by considering that the resolution of the RIO04 MDT is higher than our model, the former therefore may have current systems slightly misplaced with respect to those simulated by the ocean model.

d. Characterization of sea-level anomaly observational error variance

The error variance associated with the sea-level anomaly observation may be thought of as the sum of several contributing terms: the instrumental error variance $\boldsymbol{\sigma^2_m}$, the representativeness error variance $\boldsymbol{\sigma^2_{rep}}$, the error variance of the LHA scheme $\boldsymbol{\sigma^2_{lha}}$ (i.e. the one associated to the observation operator), and the error variance $\boldsymbol{\sigma^2_{MDT}}$ associated with the Mean Dynamic Topography. Formally, the SLA error variance $\boldsymbol{\sigma^2_{sla}}$ reads as follows:

$$\boldsymbol{\sigma^2_{sla}} = \boldsymbol{\sigma^2_m}\left(\boldsymbol{sat}\right) + \boldsymbol{\sigma^2_{rep}}\left(\boldsymbol{\lambda}, \boldsymbol{\phi}\right) + \boldsymbol{\sigma^2_{lha}}\left(\boldsymbol{\phi}\right) + \boldsymbol{\sigma^2_{MDT}}\left(\boldsymbol{\lambda}, \boldsymbol{\phi}\right). \tag{12}$$

In formulating Equation (12), we have assumed that i) the instrumental error is constant in time and depends only on the instrument itself, namely the satellite. Note that also the accuracy of the preprocessing techniques (i.e. geophysical and orbital error removals) contributes to this error source; ii) the representativeness error may vary in space but it

is temporally uncorrelated; iii) the accuracy of the observation operator (Equation (7)) is related to the degree of geostrophy. It is assumed to be dependent on the latitude only; iv) the MDT error is spatially non-uniform, as it depends on the model SSH variability.

The instrumental error has been set equal to 2 cm for TOPEX/Poseidon and 3 cm for the other satellites (see Le Traon and Ogor (1998); Cheney et al. (1994) for a discussion on the accuracy of the sea-level anomaly observations and their preprocessing techniques).

The representativeness error arises since the signal contained in the observations cannot by resolved by the analysis mesh for some small scales. It cannot be easily eliminated by the data assimilation system (Mitchell and Daley 1997), and its quantification is non-trivial. The representativeness clearly depends on the model resolution, as it vanishes by definition when all the scales are resolved by the model. Such an error may be not uniform on the domain, being larger in correspondence of strong eddy activity areas (Schiller et al. 2008). To derive the representativeness error, we have followed an approach similar to that of Oke and Sakov (2008). The cell-averaged sea-level anomaly along-track products may be thought of as the *truth* in the sense that their scales are completely resolved by the data assimilation system. The standard deviation of their difference with the observations actually assimilated after background quality-check and spatial thinning represents, therefore, the error due to the model-data scales inconsistency. Doing this, we include in the representativeness error also the contribution of the thinning, as well as the low-pass Lanczos filter. In Figure 2a, we show the contour map of the representativeness error thus calculated. Clearly, it is much larger in those areas characterized by a strong eddy activity, especially in the Gulf Stream region, in the Kuroshio Current region, and in the Agulhas Current region, as it is acknowledged that sea-level height small-scale variability closely follows the eddy kinetic energy signal (Kaplan

et al. 2004). Its globally averaged value is 32 mm. The geographical distribution of the error signal is very similar to that computed by Oke and Sakov (2008) for a global 1-degree resolution grid.

The observation operator error and its dependence on the latitude has been calculated from the differences of the SSH prognosed by the model, assumed to be the *truth*, and the SSH computed by the diagnostic formulation of Equation (7). The statistics have the property to be practically unbiased (not shown), and the accuracy is of less than 2 cm far from the Equator and less than 6 cm in correspondence of the Equator.

As our Mean Dynamic Topography was firstly derived from the long-term mean of the model SSH, its error is related to the SSH anomaly variability, and has been scaled down to account for its successive refinement through the assimilation statistics optimal interpolation, as described in the previous Section. In Figure 2b the MDT error thus derived is contoured. The variability of the SSH is larger in areas dynamically more active. The structure is therefore not very different from that of the representativeness error, though the two sources of error are conceptually different: the latter arises from the unresolved scales in the model, while the MDT error from the uncertainty in the representation of a reference ocean topography.

4. Results

In this section we present the impact of the SLA assimilation, focusing also on the impact of the choice of the Mean Dynamic Topography used within the SLA observations assimilation. The experiment which uses the MDT derived from the model and adjusted via

optimal interpolation of assimilation statistics is referred to as MDTOI; the experiment based on the RIO04 Mean Dynamic Topography is referred to as RIOMDT. The control experiment without the assimilation of SLA observations is called NOSLA. Finally, MODMDT refers to the experiment where an MDT based only on the mean SSH of a model run without assimilation is used. The error variance associated to the Mean Dynamic Topography in Equation (12) when the RIO04 MDT is used was provided by the RIO04 MDT objective analysis. In Table 1 the nomenclature and main characteristics of the experiments are summarized.

a. Impact of SLA assimilation on the analyses

The global impact of the assimilation of the SLA observations on the analyses is here shown as the bias and the standard deviation of the differences of the analysis increments, in model space, with and without assimilating altimetric data for the analyses covering the period 1993-2005. These statistics aim at not accounting for the different background fields in the two experiments by using the analysis increments only. In Figure 3 the effects of assimilating SLA is plotted as contour map of the bias within the first 100 m of depth, for both temperature and salinity. The analyses result warmer in the North-Atlantic, especially in correspondence of the Gulf Stream current, and in the Kuroshio Current region. A cooling effect is also visible in the eastern Equatorial Pacific, while more complicated patterns are present in the southern hemisphere mid-latitude band. Regarding the salinity (Figure 3b), most of the impact of the SLA observations is found in the Equatorial band, especially in the Indian Ocean, and in the Gulf Stream Region. Further, an increase in the analyzed

freshwater is found in correspondence of the Amazon river outflow, where the freshwater budget formulation in the model may be insufficient to correct the salinity fields.

The zonally-averaged standard deviations of the analysis increment differences are shown in Figure 4 as a function of the depth and latitude for the first 500 m of depth. Such a statistics represents the variability of the analysis due to the SLA observations with respect to the SLA-denial experiment. We found that most of the impact on the temperature fields is located above the thermocline at mid-latitudes (60S-20S and 20N-60N), while below the thermocline in the Tropical regions, due to the deep-peaking structure of the temperature vertical covariances in those regions. Note also that, as previously mentioned, SLA observations are not assimilated when very close to Equator (2S to 2N) and are given a larger error (i.e. a smaller weight) in the Tropical region as a consequence of the weaker validity of the "level of no-motion" assumption for regions dominated by ageostrophic circulation. This means that the assimilation of SLA in the neighboring areas induces a significant modification of the density profiles also in correspondence of the Equator itself. The impact on the salinity fields is stronger in the northern hemisphere up to 150 m of depth. For both the variables, the variability of the analysis due to the SLA assimilation vanishes gradually below about 300 m and becomes meaningless below 600 m of depth.

To study the relative impact of the sea-level anomaly observations with respect to the other in-situ observation types, we computed the Degrees of Freedom for Signal (DFS, Cardinali et al. (2004)) for each observation type and parameter. DFS summarize the informative content brought by different observation types into the analysis, and have been already applied by Montmerle et al. (2007) and Storto and Tveter (2009) to study the hierarchy of observation networks within limited area meteorological variational data assimilation sys-

tems. The procedure to compute DFS is based on a randomization technique firstly explored by Chapnik et al. (2006), to whom the reader is referred for further details. The computation was performed for the period 1993-2005 perturbing one data assimilation cycle per season. The absolute DFS (Figure 5a) represent the global impact of the different observation types on the analysis, while relative DFS (i.e. DFS divided by the number of observations, Figure 5b) quantify the ideal informative content brought by a single observation, regardless of the actual amount and geographical distribution of such observations. DFS show that, besides the large importance of the Argo network, sea-level anomaly data play a role comparable with that of the other in-situ observations, in both absolute and relative terms. TOPEX/Poseidon (T/P) exhibits the larger value for the absolute DFS, due to its longer data availability, while in relative terms T/P, ERS-1 and ERS-2 are the most informative altimetric satellite. The computation performed for the pre-Argo era only (period 1993-1998, not shown) highlights that the sensitivity of the analysis to the SLA observing network (absolute DFS) is second only to the XBT network.

b. Impact of different MDTs on verification skill scores

1) Sea-surface height

The impact of the mean dynamic topography has been firstly evaluated by comparison of assimilation statistics among different experiments. In terms of innovations bias and RMSE relative comparison, Figure 6 depicts the global background bias and RMSE timeseries for MDTOI, RIOMDT and MODMDT experiments. Clearly, the use of the MDTOI reduces the positive bias which is found in the RIOMDT experiment, and the root mean square

error decreases constantly of about 5 mm till the beginning of 2004. The RMSE decrease due to the use of MDTOI with respect to RIOMDT (not shown) is evident especially in correspondence of the Antarctic Circumpolar Current, where the difference reaches values up to -2 cm, consistently with the main differences found between the two MDTs.

From about 2003-2004 an increase in the RMSE is visible in the figure, which seems to be due to a change in the ECMWF forcing fields. In particular, an increase in the net downwards water flux as seen by the ECMWF operational analyses with respect to the ERA-40 reanalyses has led in turn to a constant increase of the globally-averaged sea-level height from 2002 onwards. Also the lack in the correct detection of freshwater budget changes contributes to the increased RMSE, whereas the OGCM used in this study cannot simulate neither the effects of variations of the steric component of sea-level nor the changes in the ocean mass due to ice melting, the latter being more and more important in relative terms in the last decade (Cazenave et al. 2009; Leuliette and Miller 2009). As additional reason for the RMSE increase, note also that since June 2003 the launch of Jason-1 caused a noticeable increase of SLA observations amount, which may cause a change in the skill score statistics.

Furthermore, the error increase due to the use of MODMDT (visible in Figure 6) which does not account for observations is non-negligible from 1997 onwards, and highlights the benefits of the MDTOI. Further to the South Hemisphere, the error increase is also relatively large (up to 1 cm, not shown) in the Gulf Stream region.

The impact of the assimilation of SLA on the SSH itself is non-trivial, as our analysis system does not correct the SSH. The use of a model-derived MDT implies also that the SSH variability rather than the absolute SSH is assimilated. Model sea-level height has been verified against independent measurements from many islandic (i.e. non-coastal) tide gauge

stations, whose data availability covered the period 1993-2005. Tide gauge observations were taken from the Research Quality Data Set of the Joint Archive for Sea Level (JASL), provided by the University of Hawaii Sea Level Center. In particular, 4 stations in the Atlantic, 2 in the Indian and 20 in the Pacific were chosen for this study according to their data availability for the period 1993-2005 and their distance from the coast. Table 3 summarizes the impact of sea-level anomaly observations. We found that the impact of the SLA assimilation on the Atlantic stations verification is not very significant. More significant is the impact on the two stations located in the Indian Ocean where the RMSE decreases of about 1 cm. Different impact have been found for the 20 Pacific Ocean stations, where 5 stations exhibited an almost neutral impact (RMSE difference less than $\pm$ 0.5 cm), 3 stations a slightly positive impact (RMSE difference between 0.5 and 1.5 cm), and 12 stations a positive impact (RMSE decrease greater than 1.5 cm). The impact of the different MDT is smaller compared to the impact of the SLA observations themselves.

2) In-situ observations

Verification against high-quality in-situ observations from the Tropical Atmosphere-Ocean (TAO) project in the Equatorial Pacific, the Research Moored Array for African-Asian-Australian Monsoon Analysis and Prediction (RAMA) in the Equatorial Indian and the Prediction and Research Moored Array in the Atlantic (PIRATA) have been conducted for the entire experimental period. These datasets represent a unique source of information about temperature and salinity variability from moorings in the Equatorial region. While temperature data are measured by means of very precise thermistors, salinities are calculated

from the conductivity and temperature values using the method of Fofonoff and Millard Jr (1985). For verification purposes, we used averaged daily values.

The root mean square error against the three observing arrays are shown in Figure 7 for the first 150 m of depth as a function of the forecast length. In the Equatorial Pacific (TAO) the impact of SLA observations is positive on the temperature regardless of the MDT used in the experiment, while for the salinity the use of the observed MDT degrades the skill scores with respect to the experiment without SLA assimilation. A similar result is found in the Indian Ocean (RAMA), although the impact of SLA on the temperature skill score is much slighter. Finally, the skill scores in the Equatorial Atlantic are worse when the RIO04 MDT is used also for temperature, while the SLA impact is neutral for temperature and salinity in the case that the MDTOI is used. In these verification statistics, we have also verified the MODMDT experiment, which for all the oceans and for both parameters exhibits a very negative impact. To summarize, the MDTOI experiment is the one that shows the most positive impact for both temperature and salinity.

Verification of the ocean heat content has been conducted against the heat content dataset of Levitus et al. (2009). They provide an estimate of the heat content in the first 700 m of depth for the period 1955-2008 by using bias-corrected and quality-checked in-situ measurements of subsurface temperature from bathythermographs (XBT and MBT), Ocean Station Data casts (OSD) and Conductivity-Temperature-Depth (CTD). The comparison was performed in terms of seasonal mean anomaly correlation. Heat content data in general fit well with model heat content in the North and Equatorial Pacific, and partly in the North Atlantic and Indian Oceans, while the comparison worsens in the South Hemisphere Extra-Tropics (where, however, there is a much smaller amount of observations). In Figure 8a the

anomaly correlation map of the heat content is contoured for the experiment MDTOI. In Figure 8b we show the difference of the anomaly correlation between the MDTOI experiment and the NOSLA experiment (i.e. positive values indicate a model-data correlation increase when SLA observations are assimilated). We find significant improvement of the model heat content caused by SLA assimilation in the North and South Pacific, in the Indian Ocean and partly in the Atlantic, while the impact of altimeter data is noisy and often negative in the high-latitude South Hemisphere. The use of the RIO04 MDT (not shown) shows in general features similar to the MDTOI impact, although one can note a slightly better impact in the Pacific and a worse impact in the Equatorial Atlantic and the band between -70 and -40 degrees of latitude, while using the MODMDT worsens the correlation scores in the North Atlantic only.

3) Surface currents

The impact of sea-level anomaly observations on the surface currents was studied through comparing monthly means of the zonal component of surface currents against the Ocean Surface Current Analysis-Real Time (OSCAR, Bonjean and Lagerloef (2002)) dataset, which derives surface currents from satellite altimeters and scatterometers (surface vector wind data). Although this dataset is not completely independent from our experiments as it is obtained also from sea-level anomaly data and adjusted by means of surface vector wind, it was proved to have good quality especially within tropical regions. Further, our assimilation scheme does not correct velocity fields and, therefore, the comparison is a proxy of how the assimilation of sea-level anomaly observations is able to induce a correct modification in the

tropical and sub-tropical surface circulation. The agreement of the model surface circulation is particularly evident in the Tropics and in the North Pacific, while weaker elsewhere (see Figure 9a for the surface zonal current correlation between OSCAR data and the experiment MDTOI). The impact of SLA data is shown as difference between the surface zonal current correlation of MDTOI experiment and that of the NOSLA experiment (Figure 9b). Largest positive impact are found within the band 30S to 30N degrees of latitude in all the Oceans. An increase of up to 0.4 in the correlation is particularly clear in the Tropical Pacific. A very similar positive impact was also found when the RIO04 MDT or the MODMDT were used (not shown), although a deterioration of the verification scores was found in close proximity of the Equator (about 3S to 3N) in both the Pacific and the Atlantic Ocean.

5. Summary and discussion

In this paper we have described a strategy for assimilating along-track satellite altimetric observations in a new global ocean three-dimensional variational data assimilation system and investigated its benefits. The 3DVAR formulation implements a First Guess at Appropriate Time algorithm with an assimilation time-window of ten days. Horizontal and vertical background-error covariances are assumed independent.

Further to all the quality-checked in-situ observations, we assimilate along-track sea-level anomaly observations from TOPEX/Poseidon, ERS-1 and -2, Geosat Follow-On, Jason-1 and Envisat provided by AVISO and treated by means of a one-dimensional low-pass Lanczos filter to filter out the signal at scales smaller than the model resolution. Along-track SLA data were assimilated via a local hydrostatic adjustment scheme, which splits the sea-level

increment, proportional to the water-column integrated density increment, into thermo- and halo- steric contributions, provided that the increment is spread into vertical profiles of temperature and salinity according to the local structure of the bivariate background-error vertical covariances.

We have built an *ad hoc* global Mean Dynamic Topography from model mean SSH, which was subsequently adjusted through assimilation statistics, whereas the MDT bias may be reasonably considered the main contributor to the observation minus first guess bias. This approach aims at assimilating the sea-level temporal and spatial variability rather than the *absolute* sea-level height, which would require a full consistence between an observations-based MDT and the model mean SSH. The large-scale features of the MDT thus computed are in general very similar to those derived from gravimetry and in-situ data, like the MDT from Rio and Hernandez (2004), except in some sensible areas, such as the Gulf Stream region and mostly in correspondence of the Antarctic Circumpolar Current, the former due to the different MDT resolution, the latter to the poor amount of in-situ observations.

We have devoted many efforts to rigorously define the observational error associated to the SLA observations, which was formally decomposed in i) instrumental error; ii) representativeness error, deducted from the differences of gridpoint-averaged SLA super-observations and raw altimetric data; ii) observation operator error, i.e. the accuracy of our local hydrostatic adjustment scheme with respect to the more accurate prognostic free-surface formulation contained in a global ocean model and iv) the MDT error computed from the SSH anomaly variability.

The experiment including the SLA observations assimilation was extensively compared with a SLA-denial experiment and with an experiment which used the RIO04 MDT. As self-

consistency proof of the LHA scheme, we have found a good improvement in the SSH fields, which are however not corrected by the 3DVAR system when SLA data are assimilated, against both sea-level anomaly data and non-coastal tide gauge stations, especially in the Pacific Ocean. For the SSH field the positive impact of assimilating SLA seems to be quite robust for all the MDTs that we used. Comparison with high-quality dataset in the Tropical regions (TAO, PIRATA and RAMA) prove that the SLA assimilation is clearly of benefit for both temperature and salinity skill scores. However, the results are quite sensitive to the MDT that we used.

We found that the definition of the Mean Dynamic Topography is crucial for the impact of the altimetric data. The results have shown that neglecting either the OGCM dynamics and SSH or the observational dataset can lead to inconsistencies between spatial variations of SLA model-equivalents and observations, which in turn affects the quality of the analyses. In particular, the use of a gravimeter-based MDT is detrimental to the Tropical Pacific and Indian salinity skill scores, and to the Tropical Atlantic temperature scores with respect to the experiment with the assimilation of in-situ observations only, while the use of the MDT that we built *ad hoc* improves the scores for both temperature and salinity in the Pacific and Indian Oceans. A model only-based MDT further deteriorates the verification skill scores for both temperature and salinity in the Tropics.

Another positive impact of the SLA observations assimilation strategies was found in the surface current verification against the OSCAR dataset: the assimilation of SLA significantly improves the tropical circulation, especially in the Pacific Ocean.

It is important to stress that with the present configuration of the OGCM, we have been able to simulate only the sea-level variations due to advective and evaporation minus

precipitation terms, and not the steric variations (due to the Boussinesq approximation in the ocean model formulation) nor the contribution of ice melting (due to the lack of estimates of continental ice melting and the use of a climatological river runoff database). The model SSH variations have much smaller variability compared to those observed by altimetric satellites. Therefore, as the ocean mass variations are not properly simulated, the success of a MDT based on the model and adjusted via observations resides on the importance of simulate spatial variations of sea-level height rather than absolute values of sea-level height.

As a final remark on future developments, we shortly plan to increase the analysis system resolution to an eddy-permitting one (0.25 degrees on both zonal and meridional directions). The relative amount of sea-level anomaly observations with respect to other in-situ observations will increase noticeably, and their impact is expected to step up accordingly. We also plan to release the assumption of spatially uncorrelated observational error for the SLA observations due to the increased resolution, which will be comparable with the altimetry correlation length-scale (Le Traon and Hernandez 1992; Le Traon and Dibarboure 1999).

Acknowledgments.

This work has been partly supported by the INGV study program "Programma Internazionale di Studi Avanzati sull'Ambiente e sul Clima", funded by the "Fondazione Cassa di Risparmio di Bologna" and partly by the European Commission-funded project MyOcean. The authors want to thank the AVISO team for the support in the use of SLA data, Dr. Simon Good (U.K. Met Office) for the support in the use of the EN3 dataset, the TAO Project Office of NOAA/PMEL for letting the authors use the TAO/RAMA/PIRATA dataset.

REFERENCES

Alessandri, A., B. S., S. Masina, A. Cherchi, S. Gualdi, A. Navarra, P. Di Pietro, and A. Carril, 2010: The INGV-CMCC Seasonal Prediction System: improved ocean initial conditions. *Mon. Wea. Rev.*, doi:10.1175/2010MWR3178.1.

Alves, J. O. S., K. Haines, and D. L. T. Anderson, 2001: Sea level assimilation experiments in the Tropical Pacific. *J. Phys. Oceanogr.*, **31**, 305–323.

Bellucci, A., S. Masina, P. Di Pietro, and A. Navarra, 2007: Using temperaturesalinity relations in a Global Ocean implementation of a multivariate data assimilation scheme. *Mon. Wea. Rev.*, **135**, 3785–3807.

Bonjean, F. and G. S. E. Lagerloef, 2002: Diagnostic model and analysis of the surface currents in the Tropical Pacific Ocean. *J. Phys. Oceanogr.*, **32**, 2938–2954.

Byrd, R. H., P. Lu, J. Nocedal, and C. Zhu, 1995: A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing*, **16**, 1190–1208.

Cardinali, C., S. Pezzulli, and E. Andersson, 2004: Influence-matrix diagnostic of a data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **130**, 2767–2786.

Cazenave, A., K. Dominh, S. Guinehut, E. Berthier, W. Llovel, G. Ramillien, M. Ablain, and G. Larnicol, 2009: Sea level budget over 2003-2008: A reevaluation from GRACE space gravimetry, satellite altimetry and Argo. *Glob. Planet. Change*, **65**, 83–88.

Chapnik, B., G. Desroziers, F. Rabier, and O. Talagrand, 2006: Diagnosis and tuning of observational error in a quasi-operational data assimilation setting. *Quart. J. Roy. Meteor. Soc.*, **132**, 543–565.

Cheney, R., L. Miller, R. Agreen, N. Doyle, and J. Lillibridge, 1994: TOPEX/Poseidon: The 2-cm solution. *J. Geophys. Res.*, **99**, 24 555–24 564.

Cooper, M. and K. Haines, 1996: Altimetric assimilation with water property conservation. *J. Geophys. Res.*, **101**, 1059–1077.

Courtier, P., 1997: Variational methods. *J. Meteor. Soc. Japan*, **75**, 211–218.

De Mey, P. and M. Benkiran, 2002: A multi-variate reduced-order optimal interpolation method and its application to the Mediterranean basin-scale circulation. *Ocean Forecasting: Conceptual basis and applications*, 281–306, n. Pinardi and J.D. Woods Eds., Springer Verlag.

Di Pietro, P. and S. Masina, 2009: The CMCC-INGV Global Ocean Data Assimilation System (CIGODAS). *Research Paper 0072*, CMCC, available via http://www.cmcc.it.

Dobricic, S., 2005: New mean dynamic topography of the Mediterranean calculated from assimilation system diagnostics. *Geophys. Res. Lett.*, **32**, L11 606.

Dobricic, S. and N. Pinardi, 2008: An oceanographic three-dimensional assimilation scheme. *Ocean Modelling*, **22**, 89–105.

Dobricic, S., N. Pinardi, M. Adani, C. Fratianni, A. Bonazzi, and V. Fernandez, 2008:

Daily oceanographic analyses by the Mediterranean basin scale assimilation system. *Ocean Science*, **3**, 149–157.

Duchon, C. E., 1979: Lanczos filtering in one and two dimensions. *J. Appl. Meteor.*, **18**, 1016–1022.

Fischer, M., M. Flügel, and M. Ji, 1997: The impact of data assimilation on ENSO simulations and predictions. *Mon. Wea. Rev.*, **125**, 819–829.

Fofonoff, N. P., 1985: Physical properties of seawater: A new salinity scale and equation of state of seawater. *J. Geophys. Res.*, **90**, 3332–3342.

Fofonoff, N. P. and R. C. Millard Jr, 1985: Algorithms for computation of fundamental properties of seawater. UNESCO Tech. Pap. Mar. Sci. 44, 53 pp pp.

Hayden, C. M. and R. J. Purser, 1995: Recursive filter objective analysis of meteorological fields: Applications to NESDIS operational processing. *J. Appl. Meteor.*, **34**, 3–15.

Ingleby, B. and M. Huddleston, 2007: Quality control of ocean temperature and salinity profiles - historical and real-time data. *J. Mar. Syst.*, **65**, 158–175.

Ivchenko, V. O., S. Danilov, D. Sidorenko, J. Schröter, M. Wenzel, and D. L. Aleynik, 2008: Steric height variability in the Northern Atlantic on seasonal and interannual scales. *J. Geophys. Res.*, **113**, C11 007.

Ji, M., R. W. Reynolds, and D. W. Behringer, 2000: Use of TOPEX/Poseidon sea level data for ocean analysis and ENSO prediction: some early results. *J. Climate*, **13**, 216–231.

Kalnay, E., 2002: *Atmospheric modeling, data assimilation and predictability*. Cambridge University Press, 341pp.

Kaplan, A., M. A. Cane, D. Chen, D. L. Witter, and R. E. Cheney, 2004: Small-scale variability and model error in Tropical Pacific sea level. *J. Geophys. Res.*, **109**, C02 001.

Le Traon, P., F. Nadal, and N. Ducet, 1998: An improved mapping method of multi-satellite altimeter data. *J. Atmos. Oceanic Technol.*, **15**, 522–534.

Le Traon, P. Y. and G. Dibarboure, 1999: Mesoscale mapping capabilities of multiple-satellite altimeter missions. *J. Atmos. Oceanic Technol.*, **16**, 1208–1223.

Le Traon, P.-Y., P. Gaspar, F. Bouyssel, and H. Makhmara, 1995: Using TOPEX/Poseidon data to enhance ERS-1 orbit. *J. Atmos. Oceanic Technol.*, **12**, 161–170.

Le Traon, P. Y. and F. Hernandez, 1992: Mapping the oceanic mesoscale circulation: Validation of satellite altimetry using surface drifters. *J. Atmos. Oceanic Technol.*, **9**, 687–698.

Le Traon, P.-Y. and F. Ogor, 1998: ERS 1/2 orbit improvement using TOPEX/Poseidon: the 2 cm challange. *J. Geophys. Res.*, **103**, 8045–8057.

Lee, T., et al., 2010: Consistency and fidelity of Indonesian-throughflow total volume transport estimated by 14 ocean data assimilation products. *Dyn. Atmos. Oceans*, **50**, 201–223.

Leuliette, E. W. and L. Miller, 2009: Closing the sea level rise budget with altimetry, Argo and GRACE. *Geophys. Res. Lett.*, **36**, L04 608.

Levitus, S., J. I. Antonov, T. P. Boyer, R. A. Locarnini, H. E. Garcia, and A. V. Mishonov,

2009: Global ocean heat content 19552008 in light of recently revealed instrumentation problems. *Geophys. Res. Lett.*, **36**, doi:10.1029/2008GL037 155.

Lorenc, A., 1981: A global three-dimensional multivariate statistical interpolation scheme. *Mon. Wea. Rev.*, **109**, 701–721.

Madec, G., P. Delecluse, M. Imbard, and C. Lévy, 1998: OPA 8.1 Ocean General Circulation Model reference manual. Note du pôle de modélisation no 11, Institut Pierre-Simon Laplace. Pp. 91.

Masina, S., P. Di Pietro, A. Storto, S. Dobricic, A. Alessandri, and A. Cherchi, 2010: Reanalysis in the Global Ocean at CMCC-INGV: examples and applications. Mercator Ocean Quarterly Newsletter 36. Available via http://www.mercator-ocean.fr.

Mitchell, H. L. and R. Daley, 1997: Discretization error and signal/error correlation in atmospheric data assimilation. II: The effect of unresolved scales. *Tellus*, **49A**, 701–708.

Montmerle, T., F. Rabier, and C. Fischer, 2007: Relative impact of polar-orbiting and geostationary satellite radiances in the ALADIN/France numerical weather prediction system. *Quart. J. Roy. Meteor. Soc.*, **133**, 655–671.

Oke, P. R. and P. Sakov, 2008: Representation error of oceanic observations for data assimilation. *J. Atmos. Oceanic Technol.*, **25**, 1004–1017.

Pascual, A., C. Boone, G. Larnicol, and P.-Y. Le Traon, 2009: On the quality of real-time altimeter gridded fields: comparison with in situ data. *J. Atmos. Oceanic Technol.*, **26**, 556–569.

Pinardi, N., et al., 2003: The Mediterranean ocean forecasting system: first phase of implementation (1998-2001). *Annales Geophysicae*, **21**, 3–20.

Reynolds, R. W. and T. M. Smith, 1994: Improved global sea surface temperature analyses using optimum interpolation. *J. Climate*, **7**, 929–948.

Rio, M. H. and F. Hernandez, 2004: A mean dynamic topography computed over the world ocean from altimetry, in-situ measurements and a geoid model. *J. Geophys. Res.*, **109**, C12 032.

Roullet, G. and G. Madec, 2000: Salt conservation, free surface, and varying levels: a new formulation for ocean general circulation models. *J. Geophys. Res.*, **105**, 23 927–23 942.

Schiller, A., P. R. Oke, G. B. Brassington, M. Entel, R. Fiedler, D. A. Griffin, and J. V. Mansbridge, 2008: Eddy-resolving ocean circulation in the Asian-Australian region inferred from an ocean reanalysis effort. *Progr. Oceanogr.*, **76**, 334–365.

Segschneider, J., D. L. T. Anderson, and T. N. Stockdale, 2000: Towards the use of altimetry for operational seasonal forecasting. *J. Climate*, **13**, 3115–3138.

Storto, A. and F. T. Tveter, 2009: Assimilating humidity pseudo-observations derived from the cloud profiling radar aboard CloudSat in ALADIN 3D-Var. *Meteor. Appl.*, **16**, 461–479.

Tapley, B. D., D. P. Chambers, S. Bettadpur, and J. C. Ries, 2003: Large scale ocean circulation from the GRACE GGM01 geoid. *Geophys. Res. Lett.*, **30**, doi:10.1029/2003GL018 622.

Uppala, S. M., et al., 2005: The ERA-40 re-analysis. *Quart. J. Roy. Meteor. Soc.*, **131**, 2961–3012.

Vidard, A., M. Balmaseda, and D. Anderson, 2009: Assimilation of altimeter data in the ECMWF Ocean Analysis System 3. *Mon. Wea. Rev.*, **137**, 1393–1408.

Vossepoel, F. C. and D. W. Behringer, 2000: Impact of sea level assimilation on salinity variability in the western equatorial Pacific. *J. Phys. Oceanogr.*, **30**, 1706–1721.

Wijffels, S. E., J. Willis, C. M. Domingues, P. Barker, N. J. White, A. Gronell, K. Ridgway, and C. J. A., 2008: Changing expendable bathythermograph fall rates and their impact on estimates of thermosteric sea level rise. *J. Climate*, **21**, 5657–5672.

Zheng, F., J. Zhu, and R.-H. Zhang, 2007: Impact of altimetry data on ENSO ensemble initializations and predictions. *Geophys. Res. Lett.*, **34**, doi:10.1029/2007GL030 451.

List of Tables

TABLE 1. List of experiments.

Experiment	Assimilated data	Used MDT
CTRL	No data assimilation	–
NOSLA	In-situ	–
MDTOI	In-situ, SLA	Mean SSH from NOSLA adjusted via OI of SLA bias
RIOMDT	In-situ, SLA	Rio and Hernandez (2004)
MODMDT	In-situ, SLA	Mean SSH from CTRL

TABLE 2. Temporal coverage of altimetric satellite observations as used in the assimilation system.

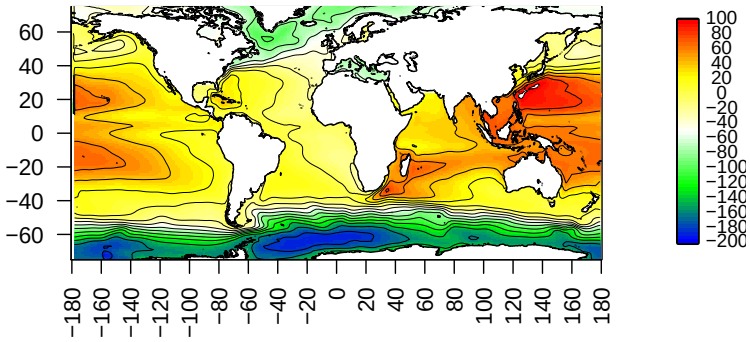
Satellite	Start period	End Period
TOPEX/Poseidon†	Sep-1992	Oct-2005
ERS-1	Oct-1992	May-1995
ERS-2	May-1995	Jun-2003
GFO	Jan-2000	Jan-2006
Jason-1	Aug-2002	Jan-2006
Envisat	Jun-2003	Jan-2006

† In new orbit from September 2002 onwards.

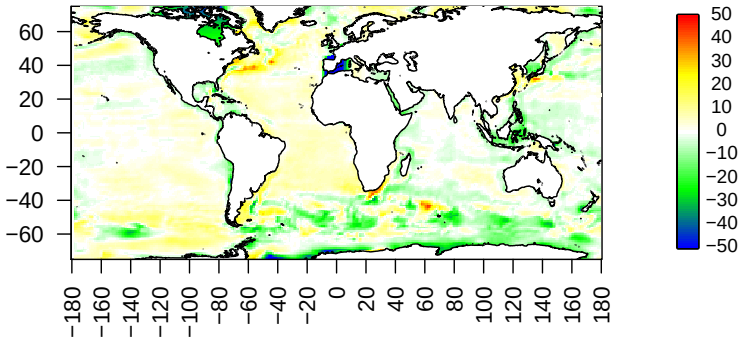
TABLE 3. Root Mean Square Error of model SSH against open-ocen tide-gauge measurements. Values are in cm.

Ocean	Number of stations	Total amount of verifying observations	NOSLA RMSE	MDTOI RMSE	RIOMDT RMSE
Atlantic	4	45150	6.16	5.84	5.89
Indian	2	27252	5.35	4.33	4.32
Pacific	20	261384	5.74	4.94	4.96

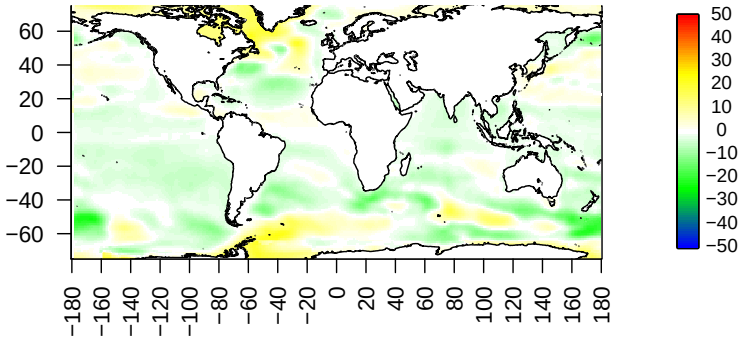
List of Figures



(a)

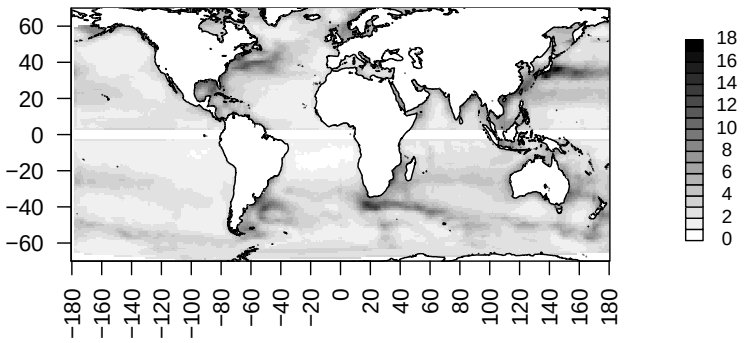


(b)

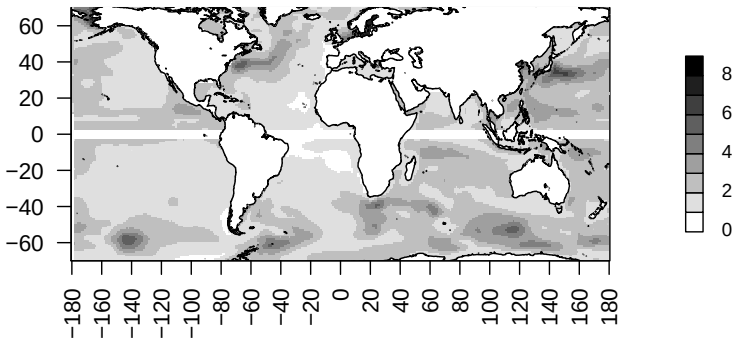


(c)

FIG. 1. Top panel: Mean Dynamic Topography computed as explained in the text and called MDTOI; contour interval is 20 cm. Difference between RIO04 MDT and MDTOI (middle) and between an MDT computed only from the model long-term SSH mean and MDTOI (bottom). Units are centimeters.

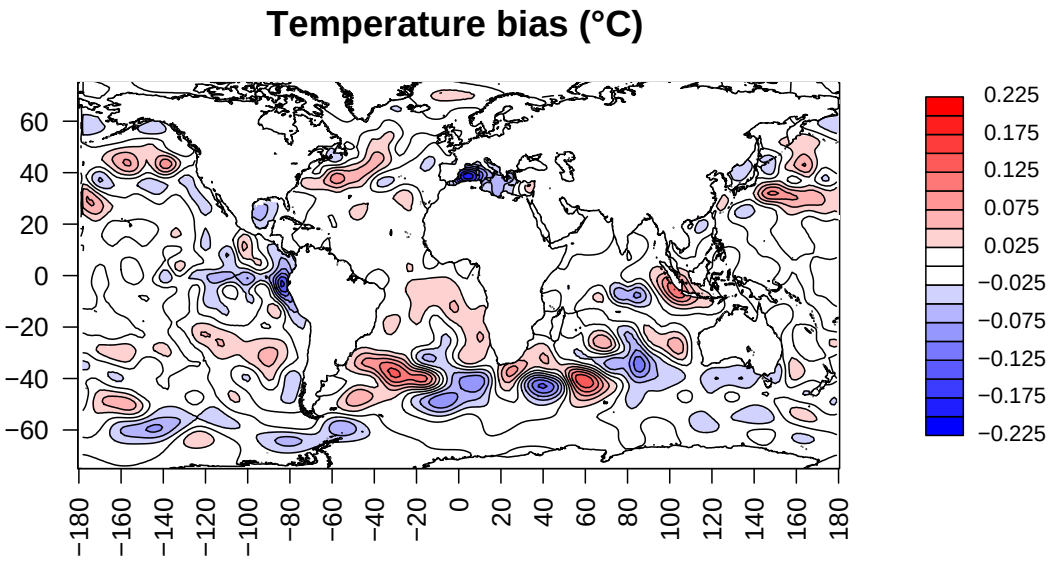


(a)

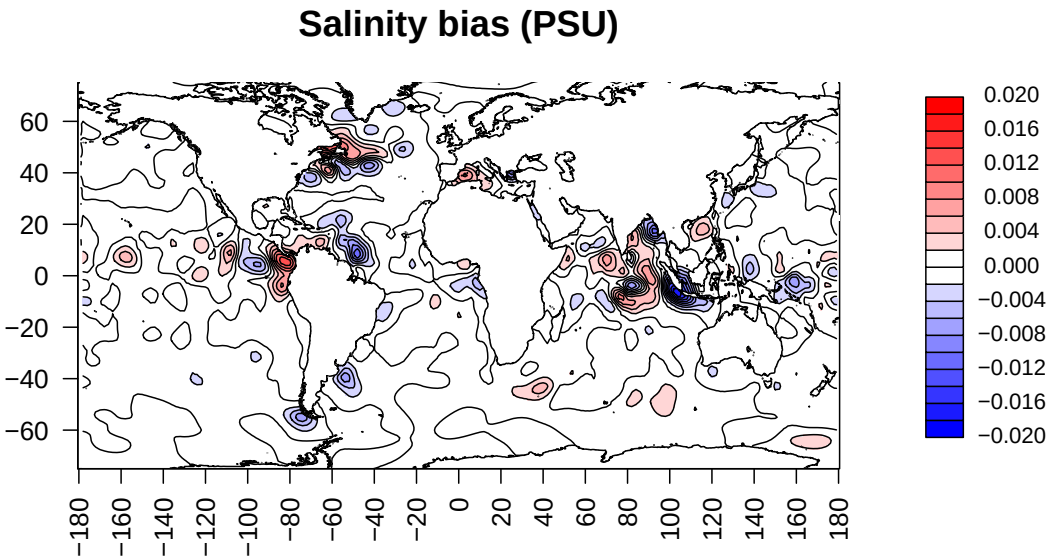


(b)

FIG. 2. Contributors to the sea-level anomaly observational error variance: representativeness error (cm) of SLA observations deducted from the differences of gridpoint-averaged SLA data and raw SLA observations (top); Mean Dynamic Topography error (cm) deducted from the SSH anomaly variability (bottom).

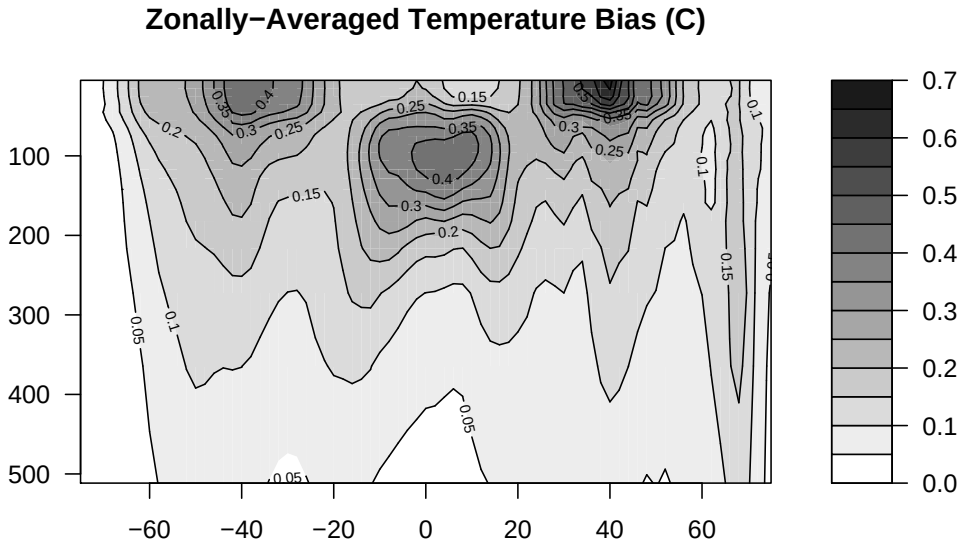


(a)

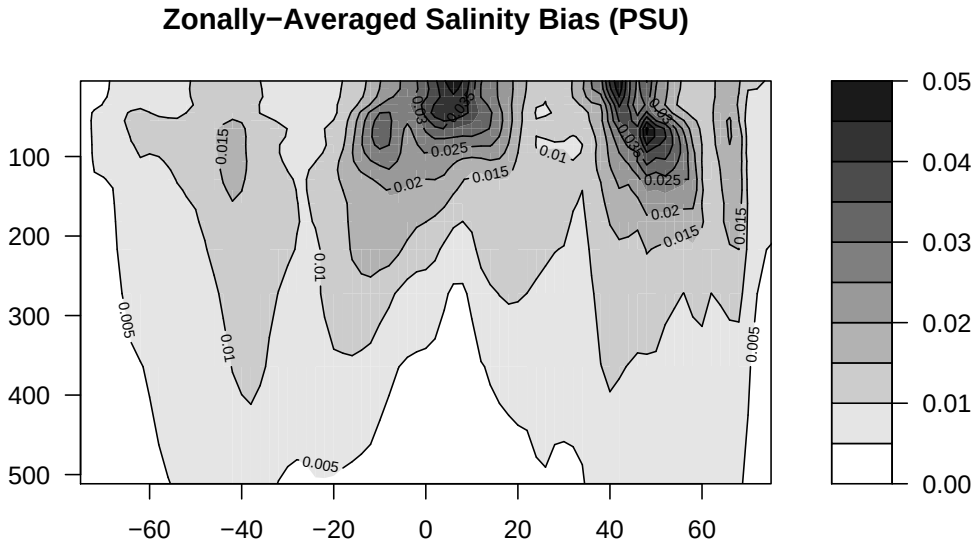


(b)

FIG. 3. Contour map of bias of the difference of temperature (top) and salinity (bottom) analysis increments with and without assimilating sea-level anomaly observations in the first 100 m of depth.

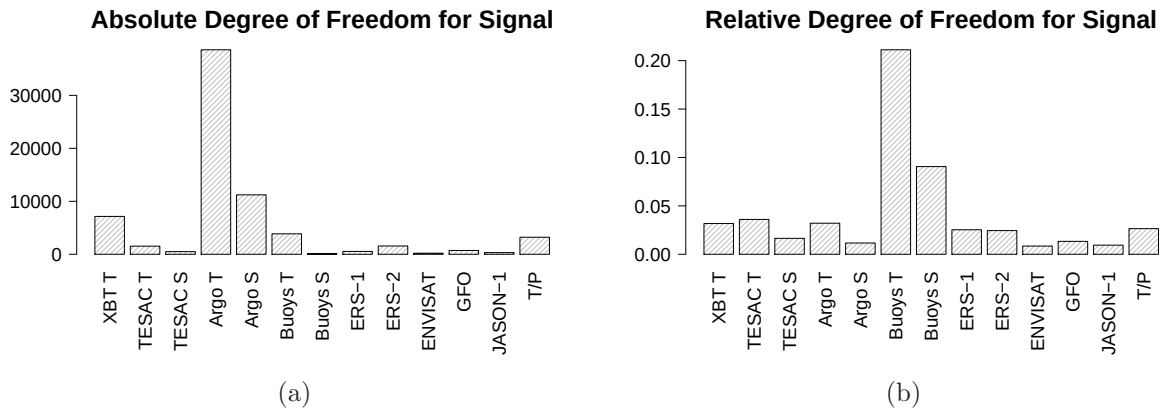


(a)



(b)

FIG. 4. Zonally-averaged standard deviation of the difference of temperature (top) and salinity (bottom) analysis increments with and without assimilating sea-level anomaly observations as a function of latitude and depth.



(a) (b)

FIG. 5. Absolute (left) and relative (right) degrees of freedom for signal (DFS), grouped by instrument and parameter (for in-situ) or satellite (for SLA).

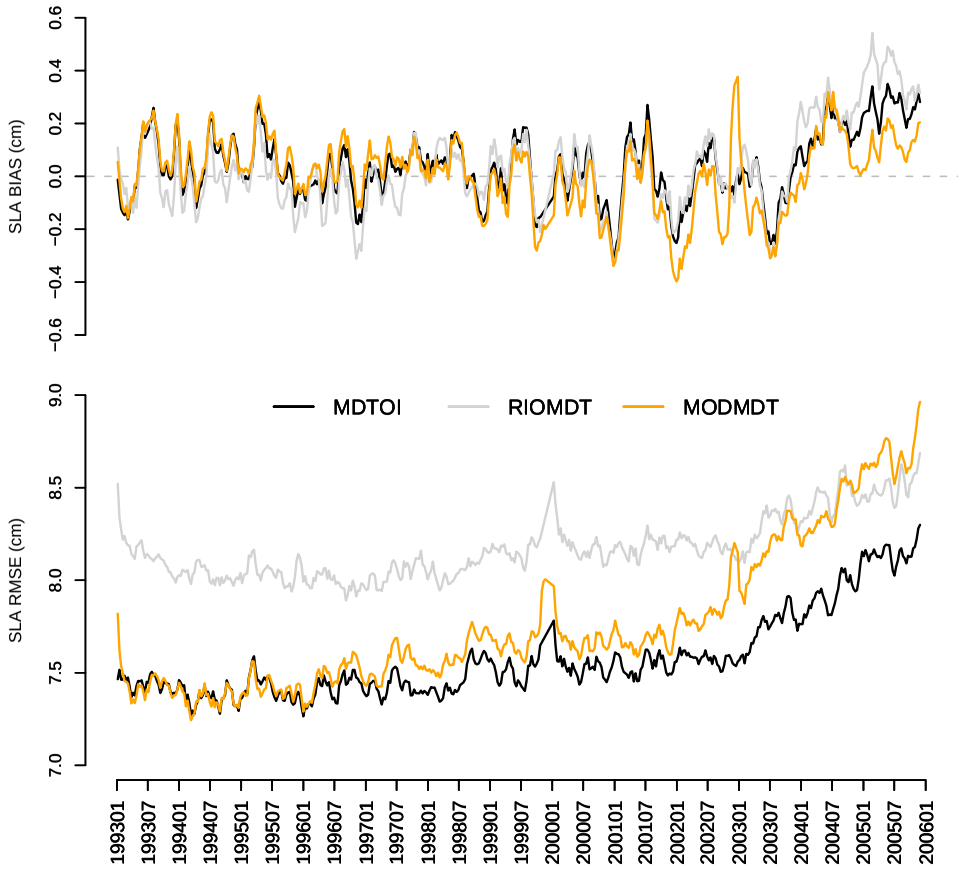


FIG. 6. Global ocean timeseries of sea-level observations bias (top) and RMSE (bottom) against model-equivalent background.

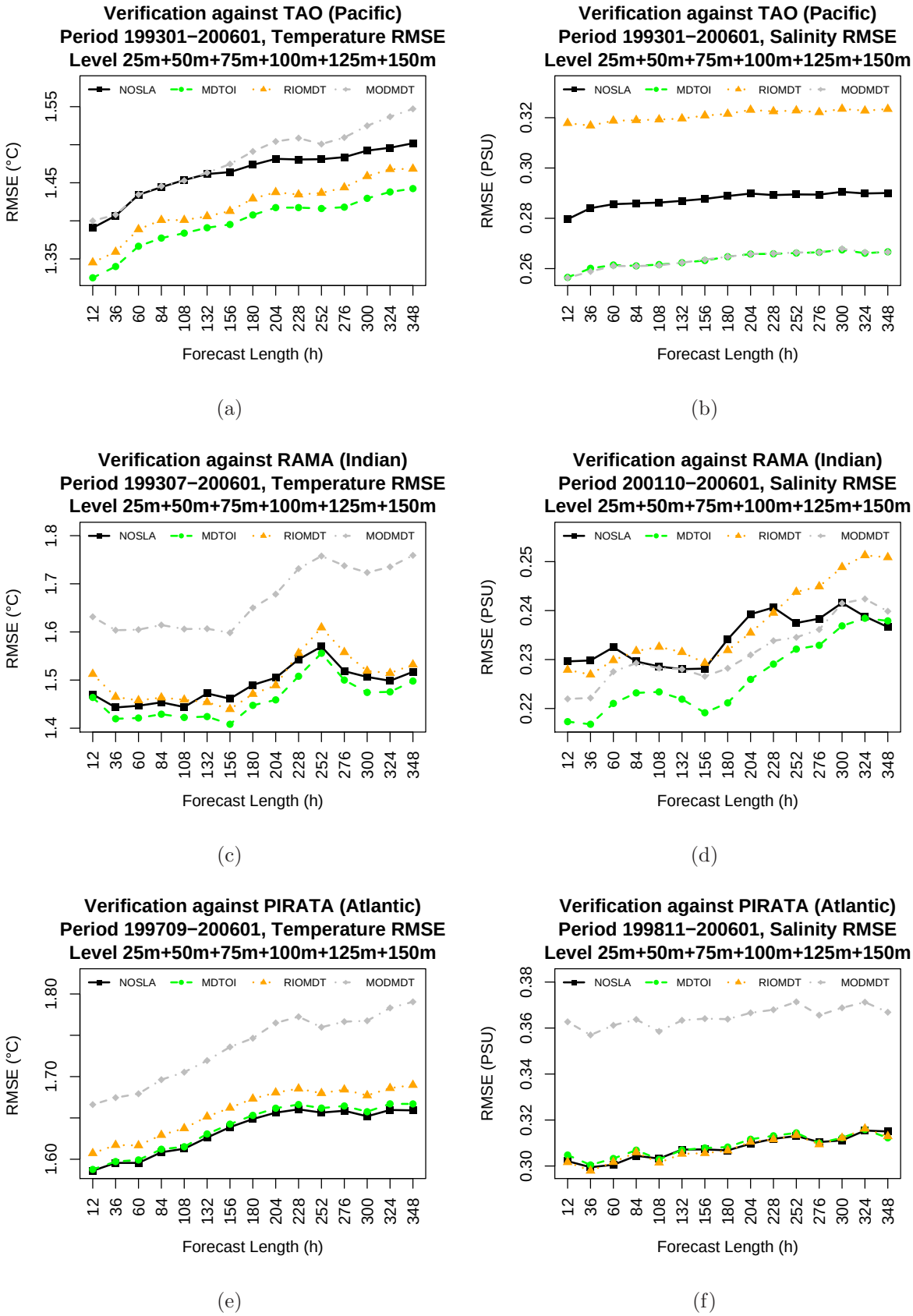
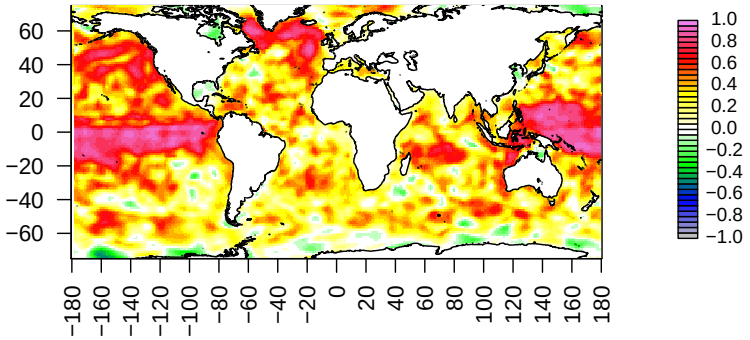
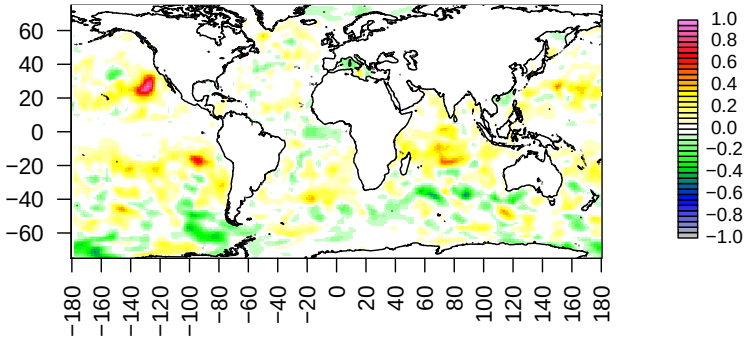


FIG. 7. Root mean square error of model temperature and salinity against TAO, RAMA and PIRATA arrays as a function of forecast length, in the first 150 m of depth.

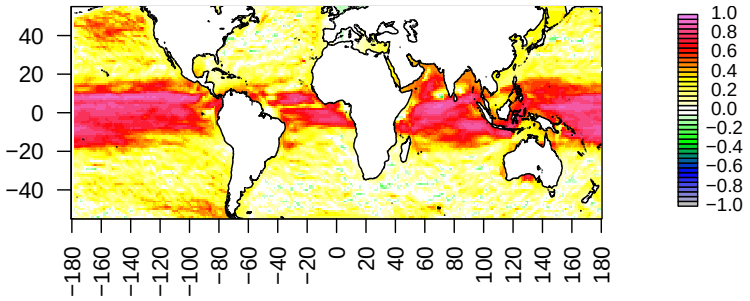


(a)

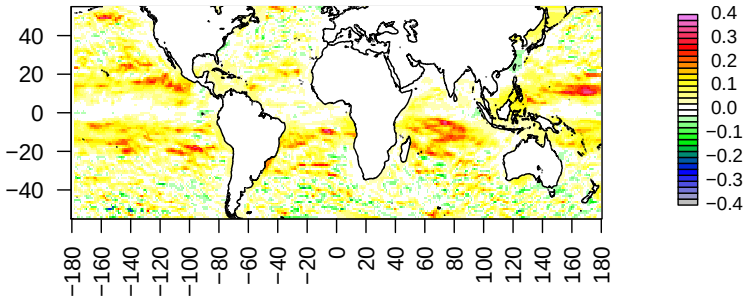


(b)

FIG. 8. Comparison of model heat content between 0 and 700 m of depth with the dataset of Levitus et al. (2009) for the period 1993-2005. Top panel: anomaly correlation of the MDTOI experiment with the dataset; bottom panel: difference of anomaly correlation between the experiments MDTOI and NOSLA: positive (negative) values indicate a better (worse) fit of the model fields with the Levitus et al. (2009) dataset.



(a)



(b)

FIG. 9. Comparison of model zonal surface current with the OSCAR dataset for the period 1993-2005. Top panel: anomaly correlation of the MDTOI experiment with the dataset; bottom panel: difference of anomaly correlation between the experiments MDTOI and NOSLA: positive (negative) values indicate a better (worse) fit of the model fields with the OSCAR dataset.